

<https://helda.helsinki.fi>

Instability of natural selection at candidate barrier loci underlying speciation in wood ants

Kulmuni, Jonna

2020-10-02

Kulmuni , J , Nouhaud , P , Pluckrose , L , Satokangas , I , Dhaygude , K & Butlin , R K 2020
, ' Instability of natural selection at candidate barrier loci underlying speciation in wood ants '
, Molecular Ecology , vol. 29 , no. 20 , pp. 3988-3999 . <https://doi.org/10.1111/mec.15606>

<http://hdl.handle.net/10138/333702>

<https://doi.org/10.1111/mec.15606>

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

DR. JONNA KULMUNI (Orcid ID : 0000-0002-8852-0793)

MR. KISHOR DHAYGUDE (Orcid ID : 0000-0003-4499-6180)

Article type : Original Article

Instability of natural selection at candidate barrier loci underlying speciation in wood ants

Short title: Instability of selection at barrier loci

Kulmuni Jonna^{1,2*}, Nouhaud Pierre¹, Pluckrose Lucy^{1#}, Satokangas Ina^{1#}, Dhaygude Kishor¹, Butlin Roger K.^{3,4}

- 1) Organismal & Evolutionary Biology Research Programme, University of Helsinki, Finland
 - 2) Tvärminne Zoological Station, University of Helsinki, J.A. Palménin tie 260, FI-10900 Hanko, Finland
 - 3) Department of Animal and Plant Sciences, Alfred Denny Building, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK
 - 4) Department of Marine Science, University of Gothenburg, S-405 30 Gothenburg, Sweden
- #) Authors contributed equally.

Corresponding author: Jonna Kulmuni, Organismal & Evolutionary Biology Research Programme, University of Helsinki, Finland, jonna.kulmuni@helsinki.fi

Abstract

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/MEC.15606](https://doi.org/10.1111/MEC.15606)

This article is protected by copyright. All rights reserved

Speciation underlies the generation of novel biodiversity. Yet, there is much to learn about how natural selection shapes genomes during speciation. Selection is assumed to act against gene flow at barrier loci, promoting reproductive isolation. However, evidence for gene flow and selection is often indirect and we know very little about the temporal stability of barrier loci. Here we utilize haplodiploidy to identify candidate male barrier loci in hybrids between two wood ant species. As ant males are haploid they are expected to reveal recessive barrier loci, which can be masked in diploid females if heterozygous. We then test for barrier stability in a sample collected ten years later and use survival analysis to provide a direct measure of natural selection acting on candidate male barrier loci. We find multiple candidate male barrier loci scattered throughout the genome. Surprisingly, a proportion of them are not stable after ten years, natural selection apparently switching from acting against to favoring introgression in the later sample. Instability of barrier effect and natural selection for introgressed alleles could be due to environment-dependent selection, emphasizing the need to consider temporal variation in the strength of natural selection and the stability of barrier effect at putative barrier loci in future speciation work.

Keywords: Speciation, Hybridization, Haplodiploid, barrier stability, genome scan, selection

Introduction

New species are formed when populations become reproductively isolated, i.e. they no longer interbreed or exchange genetic material (Mayr, 1970). However, until complete reproductive isolation has evolved gene flow may still occur between the diverging lineages. This genetic exchange is usually localized in the genome, with some regions resisting gene flow better than others. Such regions resistant to gene flow are thought to harbor “barrier loci” that can drive the divergence between lineages despite the homogenizing effect of gene flow (Ravinet et al., 2017). Barrier loci could contribute to reproductive isolation by allowing differential adaptation between diverging lineages (Ravinet et al., 2017). Alternatively, barrier loci could consist of genes

incompatible between the diverging lineages (Dobzhansky-Muller incompatibilities, DMIs (Dobzhansky, 1936; Muller, 1942)) owing to drift or divergent selection (Kulmuni and Westram 2017). Several recent studies have searched for barrier loci using genome scans and identified genomic islands of differentiation using various measures of genetic divergence (e.g. Martin et al. 2013; Soria-Carrasco et al. 2014). However, few empirical studies have confirmed that genomic islands of differentiation limit gene flow and act as barriers in the field. One strong confirmation would be observation of natural selection against introgressed alleles which maintains divergence at the barrier loci in the current population. Furthermore, the temporal stability of barrier loci is unknown in most cases. The barriers and their effect could change through time, especially when the environment changes.

We utilize hybridizing ants as a model system to map candidate male barrier loci (see Figure 1 and detailed definition below) and study the stability of their effects over time. Specifically, we use a naturally-occurring hybrid population between two mound-building wood ant species, *Formica aquilonia* and *F. polyctena* from Southern Finland (Kulmuni, Seifert, and Pamilo 2010; Kulmuni and Pamilo 2014). These species have estimated divergence within the last 500, 000 years, probably at least partly in allopatry in different glacial refugia (Goropashnaya, Fedorov, & Pamilo, 2004). The age of the hybrid population is unknown, but present-day individuals are not first-generation hybrids. Instead, the study population divides into two hybrid lineages called W and R, which likely arose from a single hybridization event as individuals from the two lineages share the same mtDNA haplotype (Kulmuni et al. 2010) (Figure 1). The two lineages were initially identified using AFLP, microsatellite and enzyme markers, which harbored several loci with alleles diagnostic to either the R or W lineage. The two lineages live in a similar habitat but mainly in different nests, have similar social structure and are indistinguishable morphologically. Even though both lineages are hybrids they show signs of reproductive isolation, as 94.5 % of successful matings are within a lineage (Kulmuni, Seifert, and Pamilo 2010; Kulmuni and Pamilo 2014) and we have never observed a reproductive adult F1 individual between the lineages. Currently we have no evidence of new alleles arriving into the population, suggesting also gene flow from parental populations into the hybrid lineages is rare.

In ants, males are haploid and born from unfertilized eggs from the mother, whereas females are diploid and arise via normal sexual reproduction. The haploid males thus allow the observation of

natural selection acting on recessive alleles (or recessive epistatic effects) that are masked in female heterozygotes (Nouhaud et al. 2020). Accordingly, our previous results suggest hybrid breakdown in males but not in females: Haploid hybrid eggs (i.e. males with introgressed alleles from the other lineage) are laid each generation, but males with specific introgressed alleles have reduced viability or die during development (Kulmuni and Pamilo 2014). Yet diploid hybrid females heterozygous with the same alleles persist in the population and are favored over more “pure” lineage individuals. This suggests recessive nuclear incompatibilities are in action in the haploid hybrid males, but these are masked in the diploid hybrid females where introgression between lineages has a benefit in heterozygotes (Kulmuni and Pamilo 2014). As a result males and females show strong allele frequency differences within a lineage at specific marker loci, females having alleles introgressed from the other lineage at frequencies as high as 40% while males lack these alleles altogether (Kulmuni, Seifert, and Pamilo 2010). Modeling suggests that these opposite selective pressures between sexes (or ploidies as males are haploid and females are diploid) can lead to either purging of incompatibilities or their long-term maintenance depending on their initial frequencies, strength of selection and recombination (Ghenu, Blanckaert, Butlin, Kulmuni, & Bank, 2018). The *Formica* system differs from the social hybridogenesis observed e.g. in *Pogonomyrmex* and *Messor* harvester ants (Helms Cahan & Keller, 2003; Romiguier, Fournier, Yek, & Keller, 2017), because in *Formica* both queens and workers arise from the same mating events and show no genetic difference in contrast to *Pogonomyrmex*.

In this study we discover and annotate candidate male barrier loci using pooled genomic sequencing and a genome scan approach (Figure 1). We define candidate male barrier loci as genomic regions of high divergence between the males of the two hybrid lineages, as these could represent recessive incompatibility loci that have remained distinct in haploid males of the two hybrid lineages despite showing admixture and signs of past hybridization and gene flow in the females. Since under the DMI model incompatible interactions among loci are thought to underlie hybrid breakdown, we further test if there is evidence of protein-protein interactions among candidate male barrier loci. Since initially identified candidate male barrier loci represent a snapshot in time, we then study the stability of their barrier effects by sampling the same population after a ten-year interval and determining whether natural selection is acting against gene flow at candidate barrier loci in males, still promoting divergence in the present-day hybrid population. For this we use SNP genotyping and survival analysis contrasting genotypes at early

developmental stage and adult stage. We expected frequencies of the introgressed alleles to decrease in males over development if the barrier effects were stable.

We find multiple candidate male barrier loci scattered throughout the genome that are fixed differently in males of the two hybrid lineages despite sharing alleles in females in 2004.

Surprisingly, a proportion of them are not stable over ten years: natural selection apparently switched from acting against to favoring introgression in males in 2014. Our results reveal the dynamic nature of hybrid populations and highlight that barrier effects at putative barrier loci may change through time. They call for more studies investigating stability of barrier effects over time.

Materials and Methods

Discovery of candidate male barrier regions with pooled genomic sequencing

We used pooled genomic sequencing to compare male and female genomes between W and R lineages and to discover the candidate male barrier loci. We defined candidate male barrier loci as SNP positions where the male genomes from the W and R hybrid lineages were alternatively fixed (i.e. $F_{ST} = 1$, see Figure 1 for experimental design). This criterion is arbitrary in the sense that many loci showing significant differentiation (but not fixed) could act as barriers and be under selection too. However, loci alternatively fixed in males, despite shared hybrid history and introgression in females are expected to experience the strongest antagonistic selection in the current population and offer the best opportunity to test for selection acting on male barrier loci. We assumed that alleles alternatively fixed in males from each lineage represented incompatible allele combinations derived from the parental populations which gave rise to the hybrid population (e.g. the alleles fixed in W males represent compatible combinations but they are incompatible with the alleles fixed in R males).

We collected the samples used for pooled sequencing from the Långholmen hybrid population (Kulmuni et al. 2010) in the year 2004. The samples were freshly frozen and kept in -20°C until genomic DNA was extracted in the year 2010 from half a body using a Qiagen kit. We sequenced four samples, each consisting of 24 individuals: R males, R females (unmated queens), W males and W females (unmated queens). These individuals were classified into the two lineages based on six to seven diagnostic microsatellite alleles (Kulmuni, Seifert, and Pamilo 2010). The sample

concentrations were checked with Qubit and pooled into the four pools in equal amounts. Due to haplodiploidy, this resulted in 24 and 48 chromosomes sampled for male and female pools, respectively. These pools represent population samples and came from two to four nests each, all of which belong to the same supercolony, where nests have hundreds of reproductive queens and relatedness approaching zero. Our previous (Kulmuni, Seifert, and Pamilo 2010; Kulmuni and Pamilo 2014) and current analyses (Supplementary Figure 1, Supplementary Figure 2) show there is no genetic differentiation between nests within a lineage, justifying pooling of individuals from different nests.

Each pool was sequenced with 100bp paired-end sequencing on its own lane in Illumina HiSeq2000 in the Institute for Molecular Medicine Finland (FIMM). This resulted in 46,106,000 to 108,204,481 total number of reads per pool. We quality trimmed reads by removing up to 20 bp that had phred score < 20 using FASTX-Toolkit. Next, we made *de novo* assemblies of each of the four samples with Soapdenovo (Li et al. 2010) trying out different kmer sizes (31, 41, 51, 61, 71) for each assembly. The R male assembly with kmer size of 41 was best in terms of completeness and quality (genome size: 222.6 Mb, 327480 contigs, average contig length: 679 bp, N50 = 1748 bp) and chosen as our reference assembly (Supplementary Table 1). This likely results in an underestimate of the true genome size as closely related *F. exsecta* (diverged ~5 Mya) has a genome size of 278 Mb. We then mapped each sample back to the R male reference assembly after removing contigs of the assembly shorter than 500bp using Bowtie2 v2.0.2 (Langmead & Salzberg 2012). Reads mapped in proper pairs and with a mapping quality superior to 20 were filtered and combined in a single mpileup file using samtools 1.4 (Li et al. 2009). Since coverage of the W male pool was low (mean = 16, s.d. = 38), overlaps between read pairs were kept. To verify that this did not influence the allele frequency estimates we tested if there was any significant difference in read counts obtained with or without filtering read pair overlaps in each pool using χ^2 tests. Over all pools, four SNPs displayed significant allele frequency difference ($P < 0.05$, for stringency no Bonferroni correction was applied) and were removed from the dataset (Supplementary Figure 3). The mpileup file was then converted in a synchronized file using Popoolation2 (Kofler, Pandey, & Schlötterer, 2011). Indels and their 5-bp flanking sequences were masked, and only bi-allelic sites displaying a minimum base quality of 20, coverage between 20 and 60 for each pool and a minor allele count of four across all pools were considered. These steps led to the identification of 166,167 SNPs for which F_{ST} estimates were computed from read

count data using Popoolation2, adjusting for differences in ploidy levels between haploid male and diploid female pools. These estimates were compared with those obtained following a more recently developed approach (Hivert, Leblois, Petit, Gautier, & Vitalis, 2018): both methods provided similar results ($r = 0.96$) and identified a similar number of SNPs differentially fixed in males (Popoolation2: 711; Poolfstat: 719, including all the loci found using Popoolation2). Popoolation2 results were more conservative (i.e, less differentially fixed SNPs) and were used for the rest of the study. For the comparison between pooled sequencing and SNP genotyping data sets (see below), allele counts were imputed from read counts using a maximum-likelihood approach (Leblois et al., 2018), considering the number of chromosomes sampled per pool.

Identification and annotation of candidate male barrier regions, protein-protein interactions and GO term Enrichment Analysis

We defined male barrier loci as SNP positions where the male genomes in the pooled sequencing samples from the W and R lineages were alternatively fixed (i.e. $F_{ST} = 1$), which led to identification of 711 SNPs. These 711 SNPs fell into 610 contigs in our assembly (cumulative size: 2.09 Mb, 0.94% of the genome), whose lengths vary between 515 – 23,815 bp (mean = 3430 bp). We created an automated pipeline to annotate genomic regions around candidate male barrier SNPs, taking advantage of the recent release of the closely-related *F. exsecta* genome (Dhaygude, Nair, Johansson, Wurm, & Sundström, 2019). First, using our assembly we extracted a sequence of 5,000 bp centered on each of the 711 candidate male barrier SNPs. This 5,000 bp cut-off was used since many of the contigs were small and we were interested in genes that harbor or are located in the near vicinity of candidate male barrier SNPs. If the contig was smaller than 5,000 bp, the full contig sequence was recovered. Since some SNPs located on the same contig, and sometimes less than 5 kb apart, 639 unique sequences were recovered and were blasted against the *F. exsecta* assembly. For each queried sequence, hits covering less than 60% of its length were filtered out and the best hit was kept per query based on e-value (over all queries, $e\text{-value} < 10^{-141}$). This allowed us to anchor 571 candidate male barrier regions on the *F. exsecta* assembly. For annotation of genes, the coordinates of these regions were extended if needed to reach 5,000 bp, so that sampling effort was equal among genomic regions. *F. exsecta* genes overlapping with best hits were collected using the GenomeIntervals Bioconductor package v1.38. Overall, 590 unique genes overlapped with our 571 candidate genomic regions. For each gene, its *F. exsecta* transcript sequence was recovered and blasted against *D. melanogaster* r6.15 proteins after filtering for the

longest protein per gene in order to use annotations from *D. melanogaster*. Alignments below 150 bp and with less than 35% identity were filtered out and the best hit was kept for each query based on e-value (over all queries, $e\text{-value} < 10^{-20}$). This pipeline recovered 317 *D. melanogaster* genes whose homologs reside in candidate male barrier regions. GO term enrichment analysis was performed for the genes in candidate male barrier regions using the Gene Ontology Consortium tool.

The hybrid breakdown model is based on epistatic interactions between genes. Molecular interactions, like protein-protein interactions, could result in epistasis. Hence, we asked if there was any evidence for molecular interactions in the form of protein-protein interactions (PPI) between the gene products in candidate male barrier regions. For this we used *D. melanogaster* proteins (317 *Drosophila melanogaster* homologs, Supplementary Table 2) and the STRING database that has information on confirmed and predicted PPIs (Szklarczyk et al., 2015). We performed 1,000 permutations to assess whether the observed PPI enrichment indicated by STRING was significantly different from what would be expected for a random set of SNPs from our dataset. For each permutation, we randomly drew 711 SNPs from the total dataset (166,167 SNPs) and annotated their flanking regions using the same automated pipeline and parameter values as presented above. The number of genes recovered per permutation varied between 267 and 368 (mean = 319). For all permutations and empirical data, PPI enrichment values were computed with STRING v10 using the STRINGdb Bioconductor package (Szklarczyk et al., 2015), setting the score threshold to 800.

Testing for barrier stability and selection by genotyping SNPs at candidate male barrier and random loci

To test for barrier stability and natural selection acting on candidate male barrier SNPs discovered from the pooled sequencing data (collected 2004) we genotyped a set of candidate and random SNPs from samples collected from the study population in 2014. These samples were used to test if significant allele frequency changes occurred between adults from years 2004 and 2014 (barrier stability) and between larval and adult stages in year 2014 (natural selection) in candidate male barrier SNPs but not in random SNPs.

Genotyping was done at individual level for a total of 196 individuals (Supplementary Table 3). Samples were randomly assigned to 96-well plates for genotyping. Primer design and genotyping were done at LGC genomics using KASP genotyping chemistry. Candidate male barrier SNPs had F_{ST} of 1 between the male genomes but F_{ST} varied between females at these loci in pooled sequencing data (collected 2004). We aimed at variable F_{ST} estimates between females at the genotyped male barrier SNPs (Figure 2). Altogether we genotyped 180 randomly chosen SNPs and 183 candidate male barrier SNPs. After removal of SNPs or individuals with more than 10% missing data, diploid males from the R lineage and three ambiguous individuals (see Supplementary Figure 2), we were left with a total of 185 genotyped individuals (for the R females; early stage: 27, late stage: 32, for the W individuals; males early stage: 32, males late stage: 32, females early stage: 31, females late stage: 31) genotyped at 300 SNPs (137 random and 163 candidate male barrier loci).

To study selection acting on candidate male barrier SNPs during development we tested for significant differences in allele counts in W males between larval and adult stages at candidate male barrier SNPs compared to the random SNPs using generalized linear mixed effects models in the R statistical software (Bates, Mächler, Bolker, & Walker, 2015). We used locus type (barrier or random) and developmental stage (larva or adult) as explanatory variables and locus_ID as a random factor. We also tested if heterozygotes for introgressed alleles at candidate loci were favored in females during development like our previous microsatellite results suggest. For this, F_{IS} values were estimated per genotyped locus and per group using the HierFstat package (GOUDET, 2005). To test if F_{IS} estimates were significantly different at candidate male barrier loci in females between larval and adult stages we used a linear mixed effects model (Bates et al., 2015) with locus type, developmental stage and group (W or R) as explanatory variables and locus_ID as a random factor.

LD could bias our results if candidate male barrier SNPs are non-independent or if some random loci are actually located close to barrier regions (i.e., hitchhiking). We took advantage of the more contiguous *F. exsecta* genome assembly and pruned the SNP genotyping data based on the location of BLAST hits in the *F. exsecta* genome (see above for BLAST parameters). This allowed us to use physical locations in the more contiguous assembly to inform us about possible linkage of genotyped SNPs, assuming collinearity between the genomes. Since the extent of

linkage disequilibrium decay is unknown, we considered three different pruning intervals for SNPs that were located on the same *F. exsecta* scaffolds: randomly discarding one of the SNPs if two SNPs were closer than 50 kb (before pruning: $N_{\text{barrier}} = 163$, $N_{\text{random}} = 137$; after pruning: $N_{\text{barrier}} = 137$, $N_{\text{random}} = 127$), 100 kb ($N_{\text{barrier}} = 121$, $N_{\text{random}} = 122$) and 500kb ($N_{\text{barrier}} = 80$, $N_{\text{random}} = 100$). We then performed the statistical tests (see below, F_{IS} and allele frequency change) with the three pruned data sets. Additionally, we computed a measure of pairwise LD (r^2) between all pairs of random and candidate male barrier SNPs for each lineage \times sex combination at the late developmental stage. Using the 50 kb-pruned data set, we first removed all SNPs which were fixed or displayed a minor allele frequency below 10% within the lineage \times sex considered. Values of r^2 were then obtained using Plink v1.90b6.10 (Chang et al., 2015) with default parameters.

Results

Multiple male barrier loci throughout the genome are candidates for hybrid male breakdown

Our pooled sequencing dataset after filtering contained 166,167 bi-allelic SNPs which we used to find candidate barrier SNPs. We defined candidate male barrier loci as SNP positions where the male genomes from the W and R hybrid lineages were alternatively fixed (i.e. $F_{\text{ST}} = 1$, see Figure 2). Note that candidate male barrier loci fall into three classes depending on the assumed direction of introgression (Figure 1b). The mean pairwise F_{ST} estimate between W and R lineages was 0.14 (s.e. = 4.66×10^{-4}) for males, and 0.07 (s.e. = 2.27×10^{-4}) for females. We found 711 SNPs (0.43 % of the total) alternatively fixed between males of the W and R lineages, but no fixed differences between the female pools (Figure 2). As expected based on our previous results, SNPs fixed differently in males were polymorphic in females and had mean F_{ST} of 0.34 between the females of the two lineages (s.e. = 5.83×10^{-3} , Figure 2). Thus, these candidate male barrier loci are differentiated also between the females of the two hybrid lineages. In addition, these loci also show significant differentiation between sexes within a lineage (Supplementary Figures 4 and 5). The alternatively fixed SNPs between the male pools mark candidate male barrier loci that are predicted to contribute to reproductive isolation and speciation through incompatible interactions in haploid males. When blasting to the more contiguous genome of closely related (5 Mya) *Formica exsecta* (Dhaygude et al., 2019) (278 Mb, 14,617 scaffolds, N50 = 997,654), our

candidate male barrier loci are distributed over multiple chromosomes (cumulative coverage of 56.2 % of the *F. exsecta* genome), but are more clustered than expected for a random set of SNPs in our data (Supplementary Figure 6). Haploid chromosome number varies between 26 and 28 in *Formica* (Rosengren, Rosengren and Söderlund 1980). Assuming evenly-sized chromosomes our candidate male barrier loci are likely to be situated in over 10 chromosomes.

Candidate male barrier loci are predicted to interact at the protein level

The genetic model for hybrid breakdown suggests inviability and sterility are caused by negative epistatic interactions between diverged genes from the parental species. One possible mechanism that can result in epistasis is molecular interaction. We asked if there is evidence for protein-protein interactions between the gene products in candidate male barrier regions. For this we used *D. melanogaster* proteins (out of 590 unique genes located within +/- 2500 bp of candidate barrier SNPs, we recovered 317 *Drosophila melanogaster* homologs, Supplementary Table 2). Our results show evidence for interactions among the genes in candidate male barrier regions (Supplementary Figure 7) as 167 genes (out of 268 genes annotated and having information in the STRING) fall into a single interaction network, with evidence for protein-protein interactions in *D. melanogaster* or other species (PPI enrichment p -value = 9.65×10^{-13} against *Drosophila* genome average). However, candidate male barrier regions are not different from the *Formica* genome average, as annotating a random set of 711 SNPs from our pooled sequencing data 1000 times showed that similar numbers of protein-protein interactions could be retrieved for random sets of SNPs. We find similar results with the functional annotations of the candidate male barrier regions. The top three terms are significantly enriched compared to the *Drosophila* genome average based on the STRING database, but we did not find any significantly enriched functions among genes in candidate male barrier regions when compared to random sets of SNPs drawn from our data. The top three biological processes characterizing our candidate male barrier genes and the *Formica* genome average are “developmental process”, “system development”, and “single-organism developmental process”. Similarly, the top three molecular functions are “ion binding”, “protein binding” and “calcium ion binding”. For a complete list of scaffolds harboring candidate male barrier SNPs and their respective annotation with gene names see Supplementary Table 2. Altogether our permutations revealed that relying on annotations from distant taxa and p -values retrieved from databases may lead to the impression of enrichment when there is none.

Two hybrid lineages are still present in the population after ten years

We tested for stability of the genetic lineages and barrier effects by genotyping individuals that were collected ten years later (2014) than the population sample used for pooled sequencing (2004). Our final dataset consisted of 100 adult individuals (recently emerged adults or late stage pupa, Supplementary Table 3), genotyped with 137 random and 163 candidate male barrier SNPs. Unfortunately, in 2014 we were able to sample only eight males from the R lineage, five of which turned out to be diploid (occasionally found from the population), thus R males were excluded from the analysis. Low occurrence of these males is consistent with their previously-estimated mortality rate of >90%. The following analyses were performed for R females, W males and W females.

The SNP genotyping confirmed that within the hybrid population, R and W lineages are still genetically distinct from one another ten years after the original population sample (Supplementary Figure 2). We found two small female larvae (discarded for the rest of the analyses), that were genetically intermediate and had unique combinations of putative parental alleles from the two lineages suggesting they were early-generation hybrids between the lineages (Supplementary Figure 2). Over all the genotyped SNPs (both random and candidate), adult allele frequencies between pooled sequencing (2004) and SNP genotyping (2014) correlated well in R females (Spearman's correlation, $\rho = 0.87$, $CI_{95} = [0.82, 0.90]$, $p < 0.001$) and in W females ($\rho = 0.91$, $CI_{95} = [0.88, 0.94]$, $p < 0.001$) (Figure 3, panels A, C). This is in line with other studies focusing on the accuracy of allele frequency estimates derived from pooled sequencing (Rellstab, Zoller, Tedder, Gugerli, & Fischer, 2013) and suggests pooled sequencing and SNP genotyping gave comparable estimates.

Proportion of the candidate male barrier loci are unstable in time

Compared to female samples from both lineages, the overall correlation between pooled sequencing (2004) and SNP genotyping (2014) was lower for the W males with a Spearman's correlation coefficient of 0.71 ($CI_{95} = [0.64, 0.76]$, $p < 0.001$, including both random and male barrier loci) (Figure 3, panel B). The lower correlation in males compared to females can be explained, at least in part, by the lower number of sampled chromosomes for (haploid) males compared to (diploid) females creating a larger error in the estimated allele frequencies. However, we also found a signal of barrier instability in the males (Figure 3, panel B and E): introgressed

(i.e. R) alleles were absent from adult W males at the 163 candidate male barrier loci in year 2004 (allele frequency of zero), but ten years later in 2014, a total of 123 loci (75.5%) harbored introgressed alleles, suggesting these SNPs mark unstable barriers. The mean frequency of introgressed alleles in 2014 in W males was 0.36, the most common introgressed allele being at a frequency of 0.81 (Figure 3, panel B). Forty candidate male barrier SNPs had remained fixed in W males. Among the 34 SNPs which were fixed in both males and females of the W lineage already in 2004, 33 were still fixed in 2014 (Figure 3 panel B, allele frequencies of (0, 0) and (1, 1)). These SNPs mark stable male barrier loci. Unfortunately, we cannot estimate how genome-wide introgression has changed during the ten years as we genotyped only a small subset of loci in 2014.

Sampling different sub-populations within a lineage in different years is unlikely to drive the signal of barrier instability and increase of introgressed allele frequencies between 2004 and 2014 in our samples of W males: we observed no population substructure within the W lineage in 2005 or 2014 (Supplementary Figures 1 and 2). Increase of introgressed alleles in males over years is not likely to be caused by errors in male allele frequency estimates either, because this should affect all candidate male barrier SNPs equally (both blue and red SNPs in Figure 3, panel B). However, this is not the case, as only one (out of 34) of the SNPs that were fixed in both males and females in 2004 is not fixed anymore in 2014 (blue SNPs in Figure 3, panel B, category I in Figure 1). In contrast, a clear majority (39 out of 46) SNPs fixed in males but polymorphic in females (red SNPs in Figure 3, panel B, category II in Figure 1) in 2004 underwent introgression in W males in 2014. Hence, the signal of barrier leakage is asymmetric involving only those loci where females previously harboured introgressed alleles but males did not (proportion test, z -value = -7.43 , $p < 10^{-5}$). Finally, drift is an unlikely explanation for the pattern of barrier instability in males, as drift is not expected to substantially increase introgressed alleles at the majority (40 out of 46) of SNPs fixed in males but polymorphic in females in 2004, as is the case (red SNPs in Figure 3, panel B). This signal of introgression at barrier loci in W males in 2014 compared to 2004 was also detected with the microsatellite panel used previously (Kulmuni and Pamilo 2014) (Fy15₂₂₂ 0 in 2004 and 0.28 in 2014; Fy3₁₉₀ 0 in 2004 and 0.45 in 2014, Supplementary Table 4). In summary, the signal of asymmetric barrier instability in W males is difficult to explain by any other reason than biological.

Selection favors introgression at candidate male barrier SNPs in W males

Next, using the samples from 2014 we tested whether candidate male barrier loci showed significant allele or genotype frequency change during development from larva ($n = 96$) to adult ($n = 100$) stage compared to the random loci. This would be consistent with natural selection acting on the candidate male barrier loci in nature. Allele frequencies were significantly different between larval and adult developmental stages in candidate male barrier SNPs compared to the random SNPs in W males (glmer, z -value = -2.98 , $p = 0.00286$), the candidate male barrier SNPs showing stronger change in frequency during development (Figure 4, red SNPs in panel B and E). The frequencies of the introgressed alleles slightly, but significantly increased from larval to adult developmental stages in W males (Figure 4, panel E), a pattern opposite to our expectations. Significant increases of introgressed alleles at candidate male barrier loci in contrast to alleles at random loci suggests selection favored the introgressed alleles in W males during development in year 2014. To account for possible effects of linkage disequilibrium (LD), the SNP data set was pruned by omitting SNPs closer than 50 kb, 100 kb or 500 kb according to their BLAST hits in the *F. exsecta* assembly. Leaving out closely situated SNPs did not affect the above results and p -values remained significant at all pruning intervals (Supplementary Figure 8, panel E). Levels of LD were comparable between candidate male barrier SNPs and random SNPs at the adult stage in 2014 in the W lineage (Supplementary Figure 9).

Candidate male barrier SNPs show significant heterozygote excess in females

Our previous microsatellite-based study showed that introgressed alleles were favored in females when heterozygous. This selection acted throughout the female lifetime. Thus, we searched for signs of similar selection by testing for heterozygote excess and for an increase in heterozygote excess between early and late stages in females using F_{IS} estimates. Indeed, candidate male barrier loci had significantly greater heterozygote excess compared to random loci already at the larval stage (GLMM, z -value = -4.404 , $p < 0.001$) and this pattern remained at the adult stage as well (GLMM, z -value = -7.133 , $p < 0.001$) in R females (Figure 5). However, candidate male barrier loci did not show a significant decrease in mean F_{IS} values (i.e. increasing excess of heterozygosity) compared to random loci during development (W females $p = 0.295$, R females $p = 0.643$). Pruning did not affect the above results (all p -values reported above from GLMMs remained significant for all pruning intervals; Supplementary Figure 8, panels A-D). We detected significantly greater pairwise LD between male barrier loci compared to random loci in adult R

females, but not in adult W individuals (Supplementary Figure 9). Greater excess heterozygosity in females at the candidate male barrier SNPs compared to random SNPs is expected based on our previous microsatellite results which showed selection favoring introgression heterozygotes in females. The fact that we cannot detect statistically significant heterozygote increase during development can result from either (i) lack of selection or (ii) selection for heterozygosity that had already acted in females at the candidate male barrier loci before the larval stage we sampled (i.e. between egg and larval stage).

Discussion

Several recent genome scan studies have used next-generation sequencing to map candidate genomic regions underlying reproductive isolation i.e. candidate barrier loci (reviewed in (Ravinet et al., 2017)). Approaches can often be based solely on comparing genomes, because observing natural selection at barrier loci, collecting data on fitness or molecular characterization of loci underlying reproductive isolation is challenging. In addition, replication of genome scan studies is sometimes done in space, but few studies are replicated in time. The stability of barrier effect is, therefore, largely unknown. We used haplodiploid ants as our model and expected the haploid males to show the strongest signals of selection against gene flow at barrier loci. Contrary to our expectations, a significant proportion of candidate male barrier loci show instability between samples collected ten years apart, selection on average favoring gene flow in males at the candidate barrier loci in the latest sample. Below we discuss reasons for the signature of barrier instability and positive selection for introgressed alleles including environment-dependent incompatibilities, formation of compatible combinations of parental alleles and recombination breaking associations between selected and hitchhiking loci. The ant system is dynamic and allows further investigations into how selection acts on hybrid genomes.

Proportion of candidate male barrier loci experience natural selection and show instability of barrier effects through time

Samples collected 10 years apart reveal that the two hybrid lineages, W and R, remain genetically distinct over time, which is partly due to the lack of successful hybrids between the lineages. In line with our previous microsatellite studies, our 2004 genome-wide data show that haploid male genomes harbor regions lacking introgression, whereas diploid females possess alleles

introgressed between lineages at these same locations. We also found an excess of heterozygosity at candidate male barrier SNPs compared to random SNPs in both female lineages. These patterns fit our earlier results (Kulmuni and Pamilo 2014), where we showed that selection favors females heterozygous for introgressed alleles. Results in the current study suggest a large proportion of the candidate male barrier loci are unstable in W males, revealing a significant difference in the genetic makeup of these males between samples collected ten years apart. A possible interpretation of significant increase of introgressed alleles in W males could be that a large proportion of the candidate male barrier loci initially identified were false positives. However, this is unlikely because, if these SNPs are just the extreme of a distribution of differentiation accumulated by mutation and drift, then their frequencies should not differ between males and females, except through sampling. Yet, as shown above, sampling error does not explain our results as it should affect all loci similarly, not a specific portion of candidate male barrier loci as we observe. In conclusion, a significant increase in the number of introgressed alleles in W males through time is unlikely to be solely caused by false positives in the candidate male barrier loci identified in 2004.

A change in the genetic make-up of W males does not necessarily reflect a significant change in the genetic composition of reproductive individuals (i.e. mated resident queens) within the population, as *Formica* queens can live over 5 years and the nests have overlapping generations. Instead, significant differences in the new males between years could also reflect differential survival of developing individuals and yearly variation in selection at male barrier loci. Indeed, we assumed the candidate male barrier loci to be linked with recessive incompatibilities, thus expecting to find evidence of negative selection during development in the haploid males (Kulmuni, Seifert, and Pamilo 2010; Kulmuni and Pamilo 2014). However, we found the opposite pattern, with introgressed alleles increasing in frequency at the candidate male barrier loci in the W males during development in the year 2014. This increase during development is consistent with natural selection favoring introgressed alleles in these males. The same pattern is observed at microsatellite loci in the year 2014; significant increase of introgressed alleles in the W males during development (Martin-Roy & Kulmuni 2019). The magnitude of allele frequency change during development observed in the present study is comparable to that observed within a generation in other systems, e.g. in stick insects for the color pattern locus (Nosil et al., 2018).

Reasons for barrier instability

Selection for, instead of against, introgressed alleles within a generation in males may be connected to barrier instability observed at these same loci. These results can be explained by at least three hypotheses, which are not mutually exclusive. First, the barriers to gene flow can break down in W males, due to recombination and subsequent formation of compatible combinations of barrier alleles, leading to relaxed selection. Under relaxed selection haplodiploid systems (and X-chromosomes) can show fluctuating allele frequencies, comparable to those observed in our data, if sexes had different allele frequencies to begin with. However, breakdown of barriers to gene flow can only explain changes observed between years (comparison between adults from 2004 and 2014) and not selection for introgressed alleles within a generation in males (early – late comparison within 2014). Also, allele frequency differences between sexes are not expected to last over multiple generations unless selection maintains them. Second, a proportion of candidate male barrier loci may have been linked with true barrier loci in 2004 due to close proximity in the genome, but due to subsequent recombination they are no longer in LD with the true barrier loci in 2014. This could lead to a situation where their frequencies are low in 2004 but not in 2014 in the males.

Again, this hypothesis can explain the change observed between 2004 and 2014 but it cannot explain the significant increase of introgressed alleles during development in males in 2014. A third possibility is that selection on candidate male barrier loci could fluctuate over time for example if the incompatibilities are environment-dependent, which could explain both within-generation and between-years effect. Environment-dependence means that the negative effects of the incompatibilities would be expressed only in certain environments. This situation is comparable to genotype \times genotype \times environment interaction, where fitness effects of a genetic variant are dependent on the genomic background it is found in, as well as the external environment (de Visser, Cooper, & Elena, 2011). Additional results based on microsatellite markers are consistent with the idea of environment-dependent selection at male barrier loci, as yearly variation in the frequency of introgressed alleles correlates with yearly spring temperature in W males across a 14-year time period (Martin-Roy & Kulmuni, 2019). Finally, a combination of breakdown of barriers and some life- stage specific selection could also explain our results. Time-series genomic data is needed to test the above hypotheses and to investigate if fluctuating temperature is a key selective pressure that creates variation in the survival of ant hybrids. These

analyses allow testing whether incompatibilities are environment-dependent, an aspect that has been to a large extent ignored in current speciation studies.

Using functional data from model organism results in overestimation of p-values for functional and PPI enrichment analyses in a non-model species

Hybrid breakdown is expected to result from deleterious epistatic interactions (i.e. intrinsic incompatibilities) between diverged alleles of the parental species (Dobzhansky, 1936; Muller, 1942). We found multiple predicted protein-protein interactions between genes located among the candidate male barrier regions, with significant *p*-values and enrichment of interactions when using the STRING database and functional information from *Drosophila*. However, similarly significant enrichment for both protein-protein interactions and functions can be retrieved for a random set of genes from our data. Our annotation of *Formica* genes is based on *D. melanogaster* homologs, which makes our pipeline biased towards conserved genes. If conserved genes are also more likely to have evidence for protein-protein interactions (as suggested by some studies (Luisi et al., 2015)), that could lead to enrichment of interactions among both candidate male barrier loci and random sets of SNPs with our data and pipeline. Our results highlight potential biases when extrapolating data from model organisms into non-model systems and call for caution when interpreting significance values for enrichment. In summary, our results suggest there is evidence for protein interactions among genes in candidate regions, but these are not significantly different from the genomic background. Functional and population studies in *Formica* are needed to test if mismatching combinations of parental alleles at candidate male barrier loci lead to hybrid inviability that could be environment-dependent, as predicted if protein-protein interactions underlie environment-dependent incompatibilities (Dobzhansky, 1936; Muller, 1942).

Conclusions

Here we bridged the gap between genome scan studies and fitness by mapping candidate male barrier loci between two recently diverged wood ant species and studying the stability of barrier effect and natural selection acting on candidate male barrier regions. For some of the candidate male barrier regions we document barrier instability between samples collected 10 years apart and selection favoring introgressed alleles in males. Future studies should investigate if this barrier leakage depends on the environment and fluctuates between years. Our results highlight the

dynamic nature of the ant system that allows investigations into genomic and molecular consequences of hybridization, areas where many questions still remain to be answered.

Acknowledgements

We thank Pekka Pamilo, Simon H. Martin and three anonymous reviewers for their valuable comments on the manuscript, Petri Kemppainen for discussions and CSC – IT Center for Science, Finland, for computational resources. JK was supported by HFSP fellowship, Marie Curie fellowship and Academy of Finland grant nro. 309580.

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, 67(1), 1–48. doi: 10.18637/jss.v067.i01
- Beresford, J., Elias, M., Pluckrose, L., Sundström, L., Butlin, R. K., Pamilo, P., & Kulmuni, J. (2017). Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown. *Molecular Ecology*, 26(15), 4013–4026. doi: 10.1111/mec.14183
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4(1), s13742-015-0047-8 . doi: 10.1186/s13742-015-0047-8
- de Visser, J. a. G. M., Cooper, T. F., & Elena, S. F. (2011). The causes of epistasis. *Proceedings of the Royal Society B: Biological Sciences*, 278(1725), 3617–3624. doi: 10.1098/rspb.2011.1537
- Dhaygude, K., Nair, A., Johansson, H., Wurm, Y., & Sundström, L. (2019). The first draft genomes of the ant *Formica exsecta*, and its *Wolbachia* endosymbiont reveal extensive gene transfer from endosymbiont to host. *BMC Genomics*, 20(301), doi: 10.1186/s12864-019-5665-6
- Dobzhansky, T. (1936). Studies on Hybrid Sterility. II. Localization of Sterility Factors in *Drosophila Pseudoobscura* Hybrids. *Genetics*, 21(2), 113–135.
- Ghenu, A.-H., Blanckaert, A., Butlin, R. K., Kulmuni, J., & Bank, C. (2018). Conflict between heterozygote advantage and hybrid incompatibility in haplodiploids (and sex chromosomes). *Molecular Ecology*. 27 (19), 3935-3949. doi: 10.1111/mec.14482
- Goropashnaya, A. V., Fedorov, V. B., & Pamilo, P. (2004). Recent speciation in the *Formica rufa*

- group ants (Hymenoptera, Formicidae): inference from mitochondrial DNA phylogeny. *Molecular Phylogenetics and Evolution*, 32(1), 198–206. doi: 10.1016/j.ympev.2003.11.016
- Goudet, J. (2005). hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5(1), 184–186. doi: 10.1111/j.1471-8286.2004.00828.x
- Helms Cahan, S., & Keller, L. (2003). Complex hybrid origin of genetic caste determination in harvester ants. *Nature*, 424(6946), 306–309. doi: 10.1038/nature01744
- Hivert, V., Leblois, R., Petit, E. J., Gautier, M., & Vitalis, R. (2018). Measuring Genetic Differentiation from Pool-seq Data. *Genetics*, 210(1), 315–330. doi: 10.1534/genetics.118.300900
- Kofler, R., Pandey, R. V., & Schlötterer, C. (2011). PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, 27(24), 3435–3436. doi: 10.1093/bioinformatics/btr589
- Kulmuni, J., & Westram, A. M. (2017). Intrinsic incompatibilities evolving as a by-product of divergent ecological selection: Considering them in empirical studies on divergence with gene flow. *Molecular Ecology*, 26(12), 3093–3103. doi: 10.1111/mec.14147
- Kulmuni, J., & Pamilo, P. (2014). Introgression in hybrid ants is favored in females but selected against in males. *Proceedings of the National Academy of Sciences of the United States of America*, 111(35), 1323045111-. doi: 10.1073/pnas.1323045111
- Kulmuni, J., Seifert, B., & Pamilo, P. (2010). Segregation distortion causes large-scale differences between male and female genomes in hybrid ants. *Proceedings of the National Academy of Sciences of the United States of America*, 107(16), 7371–7376.
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. doi: 10.1038/nmeth.1923
- Leblois, R., Gautier, M., Rohfritsch, A., Foucaud, J., Burban, C., Galan, M., ... Kerdelhué, C. (2018). Deciphering the demographic history of allochronic differentiation in the pine processionary moth *Thaumetopoea pityocampa*. *Molecular Ecology*, 27(1), 264–278. doi: 10.1111/mec.14411
- Heng Li, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools, *Bioinformatics*, 25(16):2078–2079, <https://doi.org/10.1093/bioinformatics/btp352>.
- Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., ... Wang, J. (2010). De novo assembly of

- human genomes with massively parallel short read sequencing. *Genome Research*, 20(2), 265–272. doi: 10.1101/gr.097261.109
- Lindtke, D., & Buerkle, C. A. (2015). The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. *Evolution*, 69(8), 1987–2004. doi: 10.1111/evo.12725
- Luisi, P., Alvarez-Ponce, D., Pybus, M., Fares, M. A., Bertranpetit, J., & Laayouni, H. (2015). Recent positive selection has acted on genes encoding proteins with more interactions within the whole human interactome. *Genome Biology and Evolution*, 7(4), 1141–1154. doi: 10.1093/gbe/evv055
- Martin-Roy, R., & Kulmuni, J. (2019). Temperature fluctuations between years predict temporal allele frequency variation in a hybrid ant population. *BioRxiv*, 755496. doi: 10.1101/755496
- Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., ... Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* 1817–1828. doi: 10.1101/gr.159426.113
- Mayr, E. (1970). *Populations, species, and evolution; an abridgment of Animal species and evolution*. Belknap Press of Harvard University Press.
- Muller, H. J. (1942). Isolating Mechanisms, Evolution, and Temperature. *Biol. Symp.* 6:71-125.
- Nosil, P., Villoutreix, R., Carvalho, C. F. De, Farkas, T. E., Soria-carrasco, V., Feder, J. L., ... Gompert, Z. (2018). Natural selection and the predictability of evolution in *Timema* stick insects. *Science*, 770, 765–770.
- Nouhaud, P., Blanckaert, A., Bank, C., Kulmuni, J. (2020) Understanding admixture: haplodiploidy to the rescue. *Trends in Ecology and Evolution*, 35(1), 34–42.
- Porter, A. H., & Johnson, N. A. (2010). Speciation despite Gene Flow When Developmental Pathways Evolve Published by : Society for the Study of Evolution Stable URL : <http://www.jstor.org/stable/3094654>. *Society*, 56(11), 2103–2111.
- Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., ... Westram, A. M. (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, 30(8), 1450–1477. doi: 10.1111/jeb.13047
- Reilstab, C., Zoller, S., Tedder, A., Gugerli, F., & Fischer, M. C. (2013). Validation of SNP Allele Frequencies Determined by Pooled Next-Generation Sequencing in Natural Populations of a Non-Model Plant Species. *PLoS ONE*, 8(11), e80422. doi: 10.1371/journal.pone.0080422
- Romiguier, J., Fournier, A., Yek, S. H., & Keller, L. (2017). Convergent evolution of social

hybridogenesis in *Messor* harvester ants. *Molecular Ecology*, 26(4), 1108–1117. doi: 10.1111/mec.13899

Rosengren, M., Rosengren, R., & Söderlund, V. (1980). Chromosome numbers in the genus *Formica* with special reference to the taxonomical position of *Formica uralensis* Ruzsk. and *Formica truncorum* Fabr. *Hereditas*, 92(2), 321–325. doi: 10.1111/j.1601-5223.1980.tb01715.x

Schumer, M., Powell, D. L., Delclós, P. J., Squire, M., Cui, R., & Andolfatto, P. (2017). Assortative mating and persistent reproductive isolation in hybrids. *Proceedings of the National Academy of Sciences of the United States of America*, 114 (41) 10936-10941. doi: 10.1073/pnas.1711238114

Soria-Carrasco, V., Gompert, Z., Comeault, A. a, Farkas, T. E., Parchman, T. L., Johnston, J. S., ... Nosil, P. (2014). Stick insect genomes reveal natural selection's role in parallel speciation. *Science (New York, N.Y.)*, 344(6185), 738–742. doi: 10.1126/science.1252136

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., ... von Mering, C. (2015). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, 43(D1), D447–D452. doi: 10.1093/nar/gku1003

Turner, L. M., White, M. a., Tautz, D., & Payseur, B. a. (2014). Genomic Networks of Hybrid Sterility. *PLoS Genetics*, 10(2), 18–22. doi: 10.1371/journal.pgen.1004162

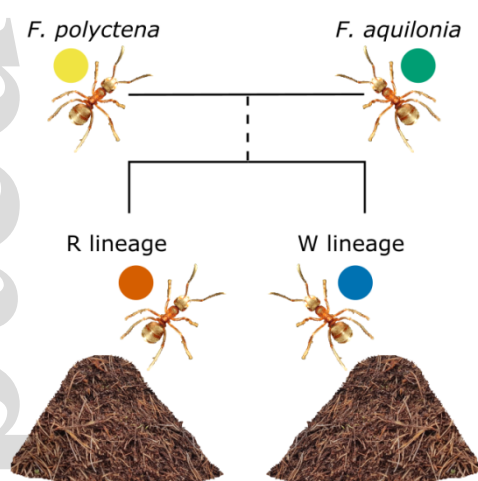
Data accessibility

Sequence data is deposited into ENA under project PRJEB39757 and genotype and other pre-processed data into Dryad under doi <https://doi.org/10.5061/dryad.w9ghx3fmn>.

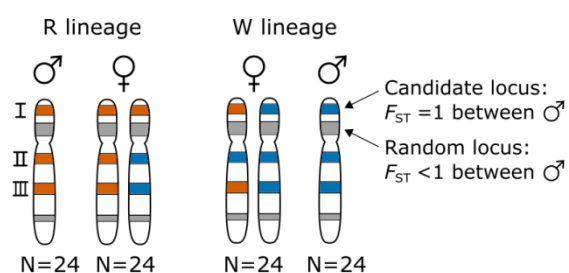
Author contributions

JK came up with the idea. JK and RB designed the study. JK, PN, LP, IS and KD analyzed the data. JK, PN, LP, IS and RB wrote the manuscript.

(a) Representation of the study system



(b) 2004, Candidate discovery in sexual adults



(c) 2014, Testing selection during development



Introgressed allele frequencies

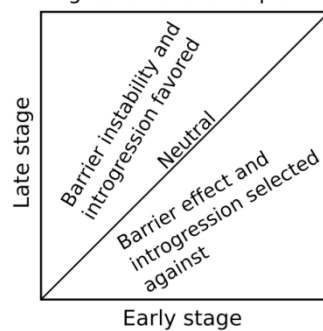


Figure 1. Overview of our approach. a) Study system, where past hybridization between *F. aquilonia* and *F. polycтена* has resulted in the coexistence of two hybrid lineages (named W and R) in different nests within the same population. b) Discovering candidate male barrier regions with pooled genomic sequencing of adults (year 2004), with a schematic representation of the male (haploid) and the female (diploid) genomes of the two lineages (simplified as a single chromosome). Red loci represent alleles originating from the R lineage and blue represent alleles originating from W lineage. Candidate male barrier loci can be divided into three classes (I= introgressed from R lineage to W females, II=introgressed from W lineage to R females and III=introgressed in both directions in females) c) Testing barrier stability and natural selection at candidate male barrier loci. Schematic representation of patterns expected in allele frequencies between larval (Early) and adult (Late) stages in 2014 under different hypotheses. Similar logic applies to comparison of allele frequencies between years 2004 and 2014.

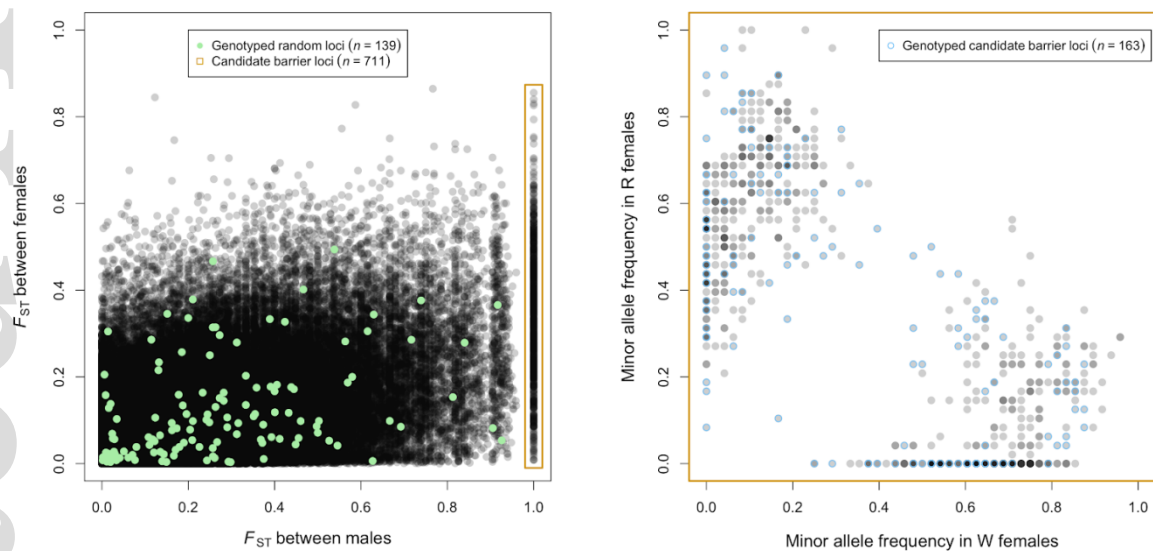


Figure 2. Differentiation between the two hybrid lineages R and W in 2004. Left panel: Genetic differentiation (F_{ST}) between the male genomes of the two hybrid lineages (x-axis) versus genetic differentiation between the female genomes (y-axis) for 166,167 SNPs called from pooled sequencing. Candidate male barrier loci ($F_{ST}=1$ between males) are in a golden box and random loci (genotyped in 2014) are in green. Differentiation between the hybrid lineages is higher between male genomes compared to female genomes (see main text). Right panel: Minor allele frequencies at candidate male barrier loci in the two females lineages (loci genotyped in 2014 circled in blue). Candidate male barrier loci differentiate females of the two lineages, with some alleles fixed in each group.

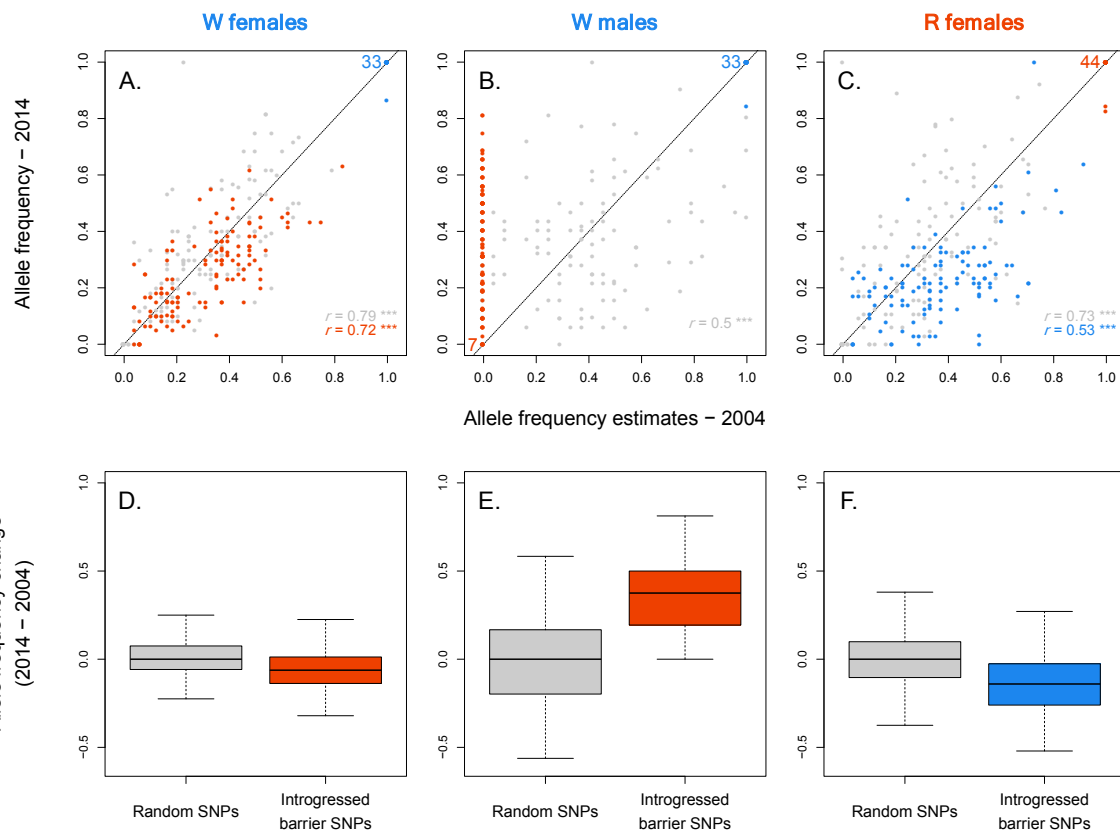


Figure 3. Allele frequency comparison between years. A-C) Allele frequencies correlate between pooled sequencing (x-axis, adults sampled in 2004) and SNP genotyping (y-axis, adults sampled in 2014) in females (lineage and sex indicated above each plot). Candidate male barrier loci are colored according to their polarization (red: R allele introgressed into W individuals, blue: W allele introgressed into R individuals). Random SNPs (in grey) are polarized by the minor allele in the pooled sequencing data. Numbers in bottom left and upper right parts indicate the number of candidate male barrier SNPs fixed in either lineage, colored according to their polarization. Spearman's correlation coefficients and their significance are indicated for each class of SNPs according to their colors (***: $p < 0.001$). D-F) Allele frequency change in candidate male barrier SNPs (red and blue) and random SNPs (grey) in each lineage and sex. Candidate male barrier SNPs show directional change in ten years' time, whereas random SNPs do not.

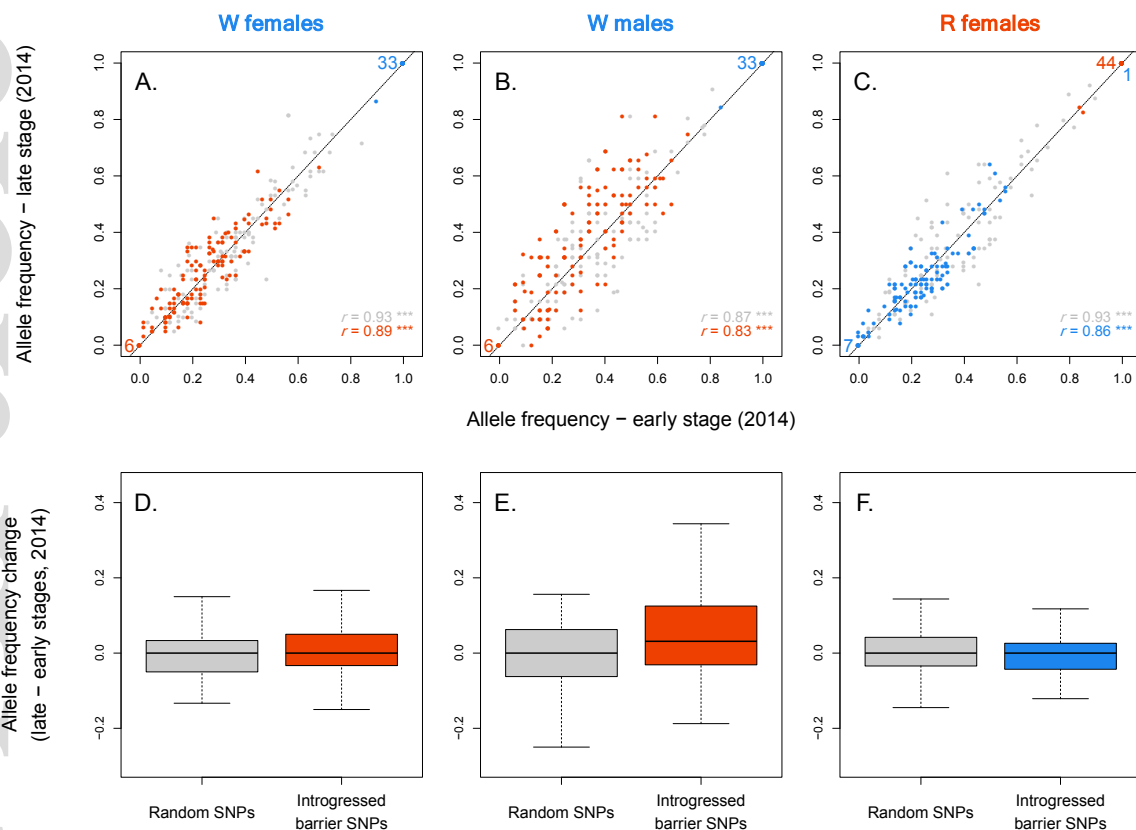


Figure 4. Allele frequency change during development. A-C) Allele frequency correlation between early (larva, x-axis) and late (adult, y-axis) developmental stages in year 2014 (lineage and sex indicated above plot) for random (grey) and candidate male barrier loci, colored according to their polarization (red: R allele introgressed into W individuals, blue: W allele introgressed into R individuals). No strong allele frequency change was observed from larva to adult in females, but frequencies in barrier SNPs in males were significantly different between larva and adult (see results), introgressed alleles increasing in frequency over development. Numbers in bottom left and upper right parts indicate the number of candidate male barrier SNPs fixed in either lineage, colored according to their polarization. Spearman's correlation coefficients and their significance are indicated for each class of SNPs according to their colors (**: $p < 0.001$). D-E) Mean allele frequency change between early (larva) and late (adult) developmental stages in random (grey) and candidate male barrier loci (red or blue) in year 2014 (lineage and sex as in A-C). Candidate male barrier SNPs have significantly stronger allele frequency change compared to random SNPs in W males ($p = 0.003$, see methods for details on the model).

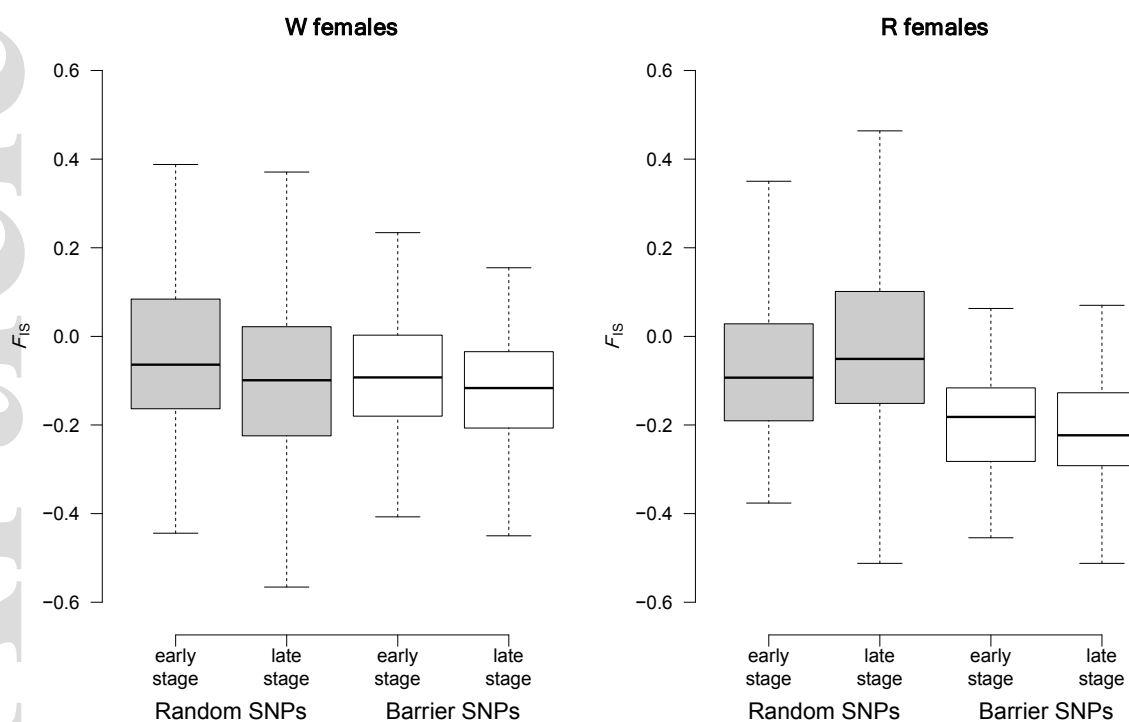


Figure 5. Mean F_{IS} estimates in early (larva) and late (adult) developmental stages in random (grey) and candidate male barrier loci (white) in the year 2014. F_{IS} estimates, particularly in male barrier SNPs, are negative showing excess heterozygosity.